AP PHOTO/CARLOS OSORIO

Measuring What Matters

A Stronger Accountability Model for Teacher Education

Edward Crowe July 2010

Center for American Progress

Measuring What Matters

A Stronger Accountability Model for Teacher Education

Edward Crowe July 2010

The Center for American Progress thanks The Bill & Melinda Gates Foundation for generously providing support for this paper.

Contents

1 Introduction and summary

2 A stronger accountability system for teacher education programs

5 Current state accountability policies and practices

5 State program approval and program review

6 National accreditation and state accountability

6 Teacher tests and program accountability

9 Federal reporting and disclosure requirements

10 Quality control and the current system

11 A stronger system of program accountability

11 Principles of a redesigned system

12 Components of a redesigned system

23 Current teacher testing fails as a quality control measure

25 Every state adopts common policies and program quality standards

28 Recommendations for states

30 Conclusion

31 References

32 Endnotes

35 About the author and acknowledgements

Introduction and summary

Our current system for holding U.S. teacher education programs accountable doesn't guarantee program quality or serve the needs of schools and students. State oversight for teacher preparation programs mostly ignores the impact of graduates on the K-12 students they teach, and it gives little attention to where graduates teach or how long they remain in the profession. There is no evidence that current state policies hold programs to high standards in order to produce teachers who can help students achieve. Moreover, every state does its own thing when it comes to program oversight—another barrier to effective quality control.

New ways of preparing teachers have been created in the last few decades in large part because they offer solutions to serious problems that many university-based teacher preparation programs appear unwilling to address. Academically strong college students as well as school districts, foundations, and policymakers are proponents of initiatives such as Teach for America, the New Teacher Project, other teaching fellows programs such as those of the Woodrow Wilson National Fellowship Foundation, and teacher residency programs.

Despite these competitive developments, however, states have done little to focus traditional preparation programs on issues like selective recruitment through high standards for entry into programs, carefully constructed and monitored clinical experiences for teacher candidates, and program evaluation focused on important outcomes.

The redesigned accountability system proposed in this paper is an effort to direct regulatory oversight to things that matter: whether or not K-12 students are learning, how well teachers have developed the classroom teaching skills to be effective with their students, a graduate's commitment to teaching as a professional career, feedback from graduates and employers, and high-quality tests of teacher knowledge and skills that are tied to classroom teaching performance and K-12 student learning.

A stronger accountability system for teacher education programs

Real quality control will hold programs responsible for how their graduates perform in classroom teaching. It will use empirically based indicators showing that students are learning from their teachers, that program graduates stay in the profession, and that they teach in the hard-to-staff schools that badly need them.

This paper argues that all states should adopt a new system of program accountability guided by these principles:

- Program accountability—and teacher preparation itself—must focus exclusively on what improves instruction and produces necessary school changes.
- State accountability for teacher preparation should be built on a set of clear signals about program quality that policymakers can understand and program faculty and institutional leaders can use.
- Signals of program quality must be empirically based, measurable indicators and should be derived from a small number of key outcomes.
- Accountability measures and their consequences for preparation programs with poor performance should be applied *equally* to all teacher preparation programs in a state, whatever the program label (traditional or alternative route) or the organization that produces new teachers.
- Full public disclosure of all program accountability findings is essential for credibility and legitimacy of state oversight policies. Clear statements, graphs, and charts devoid of jargon or evasions ought to communicate state regulators' program quality judgments.

These principles should drive development of new state accountability policies for teacher education through five essential components:

- Every state's teacher preparation program accountability system should include a teacher effectiveness measure that reports the extent to which program graduates help their K-12 students to learn.
- Classroom teaching performance of program graduates should be used by states to judge the quality of all teacher preparation programs.

- Program graduates' persistence rates in teaching should be reported for every teacher preparation program. Public disclosure of this information for up to five years post-completion will stimulate progress in addressing high teacher turnover rates by drawing attention from teacher education programs, schools, districts, and policymakers.

- Feedback surveys from preparation program graduates and from their employers should be part of state program accountability. The findings should be made public and used as a key performance indicator by all states to judge the quality of every teacher preparation program.

- A new system of teacher licensure tests should be designed and implemented for state accountability as an indicator of program quality. The number of tests now in use should be cut by more than 90 percent, and every state should adopt the same tests and the same pass rate policies.

Every state should adopt the same system of accountability indicators.

Every state should adopt the same system of accountability indicators for it to be most effective. One set of common standards would ensure that quality is defined the same way no matter where the program is located or where the graduate is employed. More than 50 versions of quality standards, policies, and accountability systems for teaching and teacher education currently exist, in contrast to engineering, nursing, accountancy, and medicine, which all have one. This paper will dig deeper into why uniformity across states is so important.

Some of the changes proposed here will take time—especially the development of high-quality tests for teacher candidates and new teachers. Even so, states can take significant steps now toward more rigorous accountability policies for teacher education programs. They can implement these four accountability measures with data systems that are already in place or on the horizon:

- Tie K-12 pupil learning outcomes to preparation program graduates and hold the programs accountable for teacher effectiveness.

- Begin to implement high-quality observational assessments of classroom teaching by supporting efforts to link these assessments to student achievement and by developing rigorous training for classroom observers to ensure reliable assessment findings.

- Employ current state data systems to track the teaching persistence rates for graduates of every program, and use the findings as a public disclosure measure.

- Implement feedback surveys of preparation program graduates and their employers using state education, labor department (or state insurance department), university, and school district data systems.

Individual states can take these steps right away. Another option is for consortia of states to work together and implement identical accountability measures and performance criteria—just as groups of states are now working on common K-12 student assessments. And finally, all states should raise passing cut-off scores on every test now in use, and they ought to make dramatic reductions in the number of redundant tests used for licensure.

Current state accountability policies and practices

If state oversight for teacher preparation programs largely ignores the impact of program graduates on the students they teach and pays little attention to where they teach and how long they remain in the profession, what is current program accountability all about? And what can we say about the effectiveness of current state policies?

An overview of state teacher education policies and practices shows that state program oversight has four main components:

- State program approval policies and practices
- A role for national program accreditation
- Licensure tests and programwide pass rates
- Policies to identify and deal with weak programs

State program approval and program review

States use program approval criteria to decide which institutions and other organizations are fit to train teachers. These policies are paired with administrative processes such as program-reporting cycles to the state, periodic reviews conducted by the relevant state agency, and campus visits by teams the state appoints.

While states may not use pupil learning results or other outcome measures to evaluate teacher preparation programs, they do put extensive time and energy into developing and administering program approval and oversight policies. These are the responsibility of the state department of education or of a "professional standards board" with jurisdiction over teacher preparation and teacher licensure. States manage oversight activities through a complex system of policy directives, teaching standards, campus visits, extensive documentation requirements, and an appellate process for programs unhappy with the results.

The details of this whole system vary greatly by state, but it's safe to say that teacher preparation program accountability is far more decentralized in the United States than in most other countries.[1]

National accreditation and state accountability

Many states complement their state regulations with national accreditation, which is a voluntary process for assuring quality control. The two national accrediting bodies for teacher education are the National Council for Accreditation of Teacher Education, or NCATE, and the Teacher Education Accreditation Council, or TEAC. NCATE accredits more than 600 programs (see www.ncate.org), while TEAC currently has just under 100 accredited programs (see www.teac.org).

Some states have partner arrangements with NCATE and conduct joint program reviews. Other states conduct reviews and make program decisions on their own. Some states substitute national accreditation for state review, while others link the two processes through formal agreements. The overlap between accreditor and state agency program oversight is confusing, and it changes regularly as states move in and out of partnerships or agreements.

State regulations and national accreditation currently pay little attention to objective data on student learning, classroom teaching, or persistence in teaching.

No compelling evidence suggests that program accreditation by NCATE or TEAC leads to (or takes account of) positive academic outcomes for students taught by graduates of accredited programs. In fact, the national accreditors for teacher education do not use empirical data on teaching and learning outcomes to make judgments about program quality. Nor is there any reason to believe that teachers who complete an accredited preparation program are more likely to demonstrate high-quality classroom teaching performance than those trained elsewhere.[2]

Teacher tests and program accountability

State regulations and national accreditation currently pay little attention to objective data on student learning, classroom teaching, or persistence in teaching. As a result, they provide little quality control for the public, schools, or parents and students.

Thanks to No Child Left Behind, however, most states now require teacher candidates and new program graduates to pass a battery of tests in order to be licensed.

Preparation program pass rates for all graduates are used as a component of state accountability in 32 states. These states typically require that a certain proportion of graduates pass the tests (80 percent in New York state, for example), and they employ the pass rate as a measure of program quality.[3]

The public can access pass rates on teacher tests for each state and for every program in the state through federal reporting requirements Congress established in 1998. Does this mean the public and policymakers should regard teacher test results as a useful tool that captures important information about preparation program quality?

Not quite. An analysis of the tests and their use across the states shows that current teacher testing has very little value as an accountability mechanism. An analysis of the tests and their use across the states shows that although they flag the very weakest programs and the weakest teacher candidates, many states work hard to avoid imposing any consequences on preparation programs no matter how poorly their candidates and graduates perform on licensing examinations.

The teacher tests now in use have many problems. One of these problems is that they don't indicate how well teachers will do in the classroom. The National Research Council reported in 2001 that teacher licensure tests "are not constructed to predict the degree of teaching success a beginning teacher will demonstrate." Two leading scholars, Suzanne Wilson and Peter Youngs, concur with this judgment and note "little evidence that … a relationship exists between teachers' scores on such tests and their teaching success." [4]

Another weakness of the tests is that they don't directly measure what teachers do in the classroom. Many teacher tests essentially measure knowledge and skills at levels more appropriate to what eighth graders are expected to know and be able to do. Those who manage to pass these basic skills tests "know how to read, write, and do basic mathematics," according to the National Research Council.[5]

Advocates for the current tests insist they are adequate quality control checks for individuals and programs. The real story is that passing scores (or "cut scores") are set low enough in many states to guarantee that nearly every graduate will pass. Case in point: Ninety-six percent of U.S. program completers (teacher education jargon for graduates) passed all state tests in 2006.[6] But as the next section indicates, states work hard to avoid making judgments about

the quality of the programs whose graduates are being tested—whatever the test results may be. An April 2010 study comparing U.S. elementary and middle school math teachers to teachers in other countries demonstrates anew the inadequacy of current basic skills and content knowledge tests as screens even for minimal teacher quality.[7]

In addition to problems with predictive validity and pass rate performance standards, states have created a crazy quilt of basic skills, content knowledge, and teaching skills assessments that add up to 1,100 different tests. Secretary of Education Arne Duncan's 2009 teacher quality report to Congress listed about 160 basic skills tests, over 100 different tests of professional knowledge, and more than 800 different content knowledge tests. Readers with a good imagination may be able to invent 800 academic disciplines or K-12 subject areas to match the number of tests, but it's not easy to find a link between this mish-mash and the everyday teaching and learning needs of our schools. The testing system for teachers is so complex that a single state can have as many as 85 different tests.[8]

The entirely predictable result of the program accountability system's testing component is to send confusing and misleading signals about what is being measured and reported. It is common practice for states to set different passing scores on the same test, and it is not unusual for a single state to adjust the cut score on a given test from one year to the next. The way current teacher testing has been designed and implemented undercuts the legitimacy of accountability itself.

But the real problem isn't testing. The pressing issue for credible and effective accountability is the tests themselves and how they are used. Better tests—linked to important teaching knowledge and learning outcomes, and validated by independent studies with transparent findings—would be important and useful accountability mechanisms for states and for programs.[9] A battery of high-quality tests of teacher knowledge, skills, and abilities could be deployed by every state, with identical passing criteria. This would establish a standard and easily understood framework for program accountability. We already do this in medicine, nursing, engineering, and accountancy without breaching federal principles, infringing on state autonomy, or traducing local values.

This paper will say more about how to get from here to there.

Federal reporting and disclosure requirements

The fourth element of state accountability for teacher education programs comes from a set of federal rules developed after Congress passed Title II of the Higher Education Amendments, or HEA, in 1998. These rules were intended to serve as a report card on teacher preparation programs in the United States and call for:

- Uniform annual reports to the state by each program and disclosure of certain information—such as licensure test pass rates—to students and to the public.

- Reports from each state to the U.S. Department of Education, compiling program-by-program data from the state and adding additional pieces of information.

- An annual report on teacher quality to Congress from the secretary of education. Title II also requires each state to establish a set of criteria to determine whether or not a program is "low performing." Federal rules require annual reports to disclose the criteria themselves. States also must make public a list of programs found to be low performing.[10]

Although they are available to policymakers and the public (at https://title2.ed.gov/View.asp), the resulting program-level and state reports are quite large and difficult to understand. Each report is the equivalent of hundreds of printed pages.

Published program pass rates were supposed to be the critical element in the federal reporting system. They were intended to shine a light on programs whose graduates were not well-enough prepared to pass a minimum competency test to become teachers. But shortly after the report card structure was established, a significant number of institutions and state agencies joined with the teacher education professional associations—the American Association of Colleges for Teacher Education, or AACTE, as well as NCATE—to work out a way to beat the reporting system. The trick they devised was requiring teacher candidates to pass all required teacher tests before being allowed to graduate. This allows programs to report 100 percent pass rates on the teacher tests. They do not have to disclose the percentage of candidates who failed one or more tests.

Quality control and the current system

Another way to gauge the effectiveness of current state oversight is to ask whether programs that fail to do a good job preparing teachers are closed down by the state or put on warning status. It is highly likely that there are programs in many states that are not up to par and should be closed down. But the record here is very clear: Less than 2 percent of all teacher education programs in the United States have been flagged as low performing by the state in which they operate since Congress required each state to develop and implement a set of criteria to identify low-performing programs in 1998.

The question is whether it makes sense for each state to have a unique set of accountability standards for teacher education.

It's important to note here that while every state is required to have policies in place to flag low-performing programs, only a handful of states have ever taken the step of labeling even one program low performing. For instance, 31 programs were identified as "at-risk or low-performing" in 2006, up from 17 programs in 2005 and 11 in 2002—out of 1,170 teacher education programs in the country.

There's a good reason for this low number: States found a way to do as little as possible with the "low-performing program" criteria. To begin with, the federal rules allowed each state to establish its own criteria to determine whether a program was low performing. States ensured that few programs would be flagged as low performing by setting this bar as low as possible.

It's little wonder, then, that one comprehensive analysis of accountability policies concluded that, "neither teacher certification nor teacher education program accreditation have inspired much public confidence." This is still the case today.[11]

There's one last issue to include in the mix as we consider how to build a more effective accountability structure. To the reader, the most striking feature of the current "system" must be its sheer complexity, including the variation from state to state in oversight policies and practices. It is safe to say that no two states are alike in their systems for accountability even if the results are similar across most of them—that is, little or no real accountability in practice.

The question is whether it makes sense for each state to have a unique set of accountability standards for teacher education. This paper argues that states have compelling reasons to work together to establish common standards, policies, and practices—including high-quality common licensure tests—to greatly improve the likelihood that every student in every school has an effective teacher.

A stronger system of program accountability

Teacher education program quality assurance policies are not doing the job. States now rely on state program approval policies and practices, national program accreditation, licensure tests and program-wide pass rates, and flimsy policies to identify and deal with weak programs. For reasons discussed above, these approaches to program oversight have little to do with outcomes that matter for students, teachers, or schools. What's more, the public doesn't understand the current accountability process, and policymakers have little confidence in the system's ability to produce strong teachers and protect the public from weak programs.

As a result, state accountability policies do not help teacher preparation programs reliably produce the teachers we need. A new and better preparation program accountability system can and should be adopted by all states and guided by the principles described below. Program accountability must focus exclusively on the factors that improve instruction, and state accountability for teacher preparation should be built on a set of clear signals about program quality. Policymakers ought to be able to understand these signals, and they ought to be useful for program faculty and program leaders.

The system's credibility and usefulness would also be strengthened if accountability measures and their consequences apply equally to all programs in a state, whether the program is "traditional" or "alternative," and no matter which organization is responsible for the preparation program.

Principles of a redesigned system

Focus accountability on important outcomes

Key "drivers" of teaching and learning outcomes must be the basis for a new system of program accountability. The Wallace Foundation recently argued that assessment systems should be tightly focused "on the most potent behaviors that can promote better learning outcomes, rather than the peripheral concerns of daily

management" in its analysis of leadership and assessment practices for school leaders.[12] This is a reasonable bedrock principle for preparation program accountability.

A stronger state accountability system should use "potent" indicators that students are learning from their teachers, that teachers are effective in the classroom as determined by objective measures, and that program graduates stay in the profession and teach in hard-to-staff schools—instead of concentrating solely on inputs or bogging down in the nuances of complex processes like program review or the hundreds of teacher tests now in use.

Key "drivers" of teaching and learning outcomes must be the basis for a new system of program accountability.

Clear signals using solid data

A new accountability system should also communicate clear signals to those who need to know whether or not a preparation program is doing a good job. This second principle is relevant because few people have confidence in current state program accountability practices. In fact, the muddled current system undermines the legitimacy of accountability itself.

Meeting these two principles requires empirical measures to be the building blocks of an effective accountability system, and these measures have to meet standards of quality and rigor to inspire confidence. Data collection, indicator calculations, and reporting practices must be transparent. This is important for professionals on the inside to feel confident that all programs are being judged by the same rules, policymakers who need assurance that the program quality indicators contain information about important outcomes, students deciding where to enroll in a preparation program, and public schools trying to make informed hiring decisions that have predictably positive results for their students. The current accountability system for teacher education fails all these tests.

Now let's look at the components of a stronger system and how these principles should be carried out in practice.

Components of a redesigned system

Every state's teacher preparation program accountability system should include a measure of teacher effectiveness that reports the extent to which program graduates help their K-12 students to learn

High-quality instruction is the main driver for student achievement, so it makes sense that teacher effectiveness measures ought to be part of preparation program accountability policies in the states. Today, however, only three states (Louisiana, Florida, and Texas) incorporate teacher effectiveness into preparation program oversight—though Delaware and Tennessee will build teacher effectiveness data into their accountability policies as they proposed in their winning Race to the Top applications. Other states are proposing such systems in their Race to the Top applications as well.

Louisiana uses value-added analyses of student academic performance to make decisions about the quality of every public or private "traditional" or "alternate route" teacher education program in the state.[13] And Florida recently began measuring and ranking its teacher education programs based on whether K-12 pupils taught by program graduates demonstrate learning gains in the classroom (More on Florida's efforts here: http://tinyurl.com/yjwd8md). Finally, Texas announced a program accountability strategy in 2009 along the same lines as Louisiana and Florida (see http://tinyurl.com/yz9jmfg). [14]

Louisiana has the longest track record with program accountability linked to teacher effectiveness. It developed a state teacher preparation program accountability system that includes graduates' impact on the pupils they teach. Value-added findings about program graduates are a major component of its accountability system. It's still unclear how Texas will implement its approach linking student achievement and program accountability. And in Florida, the standard for judgment is the proportion of program graduates who "had 50 percent or more of their students make a year's worth of progress."

Putting it into practice

There will be issues implementing teacher effectiveness measures for program accountability at scale. Improved state data systems are needed to link teacher and student data, and effective confidentiality and privacy policies are crucial. Analysis of K-12 test data also must be done carefully using appropriate statistical models and with rigorous attention to data quality issues.[15]

Further, many program graduates will be teaching in grades and subject areas that are not tested by the states. And we know that state tests can and should be built on more robust measures of learning outcomes.

These are real issues that states will have to grapple with as they move in the direction of adopting pupil achievement as a program quality indicator.

Encouragingly, value-added analyses and "growth model" calculations of student learning are proliferating as states and districts seek better ways of gauging student outcomes. Moreover, greater use of these strategies has already stimulated intense efforts to improve the tests that serve as dependent variables. Further work to refine and strengthen value-added and similar analytical methods will continue, making this teacher effectiveness measure a feasible program quality indicator in every state.[16]

States should begin now to develop teacher effectiveness as a program quality indicator and solve these problems as they go along while not underestimating the above challenges. Most states can already link student and teacher data in their K-12 system, but they are not yet able to tie practicing teachers back to their preparation program and will need to work this out. Waiting to move until every problem is resolved to universal satisfaction, however, is not responsible or realistic—especially given the pressing and unmet challenge of helping children succeed academically through the provision of highly effective teachers.

Even when teacher preparation programs are able to measure teacher effectiveness, they need to figure out how teachers get these results.

States should use the classroom teaching performance of program graduates to judge the quality of all teacher preparation programs

Even when teacher preparation programs are able to measure teacher effectiveness, they need to figure out how teachers get these results. This is important information for understanding the impact of individual teachers, but it's also important for finding out whether preparation programs are producing new teachers with the knowledge and skills to help students learn.

Classroom observation and assessment of on-the-job teaching should be used as a key measure of preparation program quality since no single measure—no matter how powerful the findings—is enough to gauge all the relevant components of teaching quality or program effectiveness. At the very least, program graduates should have acquired knowledge and experience with core teaching practices by the time they complete a program.

Laura Goe and her colleagues suggest that teacher evaluations be based on a variety of measures of student achievement and teacher practice.[17] From other work it now appears that data on new teacher job performance may be more important to estimating a teacher's impact on student learning than anything available to schools or districts at the recruitment stage of the hiring process. Other scholar-

ship suggests strongly that classroom teaching performance can be measured in reliable and valid ways.[18]

Systematically employed for program accountability, classroom assessment of teaching practices would ensure that graduates have the skills and abilities to help students learn. What's more, data generated about all new teachers from a graduation cohort—the class of 2010, for example—would help the program itself to identify the knowledge and skill sets that are making a difference in their graduates' classrooms. And classroom assessment results can highlight areas for individual teacher improvement and form the basis for targeted professional development. For preparation programs offering induction support to new graduates, teaching assessment results also would flag areas where continued development of teaching skills would improve a teacher's effectiveness in the classroom.

Data for all program graduates can work as a summative evaluation measure for program quality, and serve as a formative assessment tool for individual graduates. Findings also would be useful information for program revisions to benefit future cohorts of teacher candidates.

Putting it into practice

Implementing a program quality indicator based on classroom teaching performance would be a major step forward in providing evidence about teacher preparation that is relevant to student learning. A system of quality classroom observation supports fair judgments about preparation programs if it has sufficient rigor to produce reliable and valid findings for individual teachers and groups of teachers.[19]

Not all classroom observation or teaching assessment protocols are the same, however. Few now in use by teacher education programs meet even minimal standards of rigor. Needed for the task, according to Robert Pianta and Bridget Hamre, are "validated, standardized observational assessments of teachers' classroom instruction and interactions."[20] Fortunately, a growing number of high-quality instruments are now available for classroom observation and assessment, with plenty of evidence that these observation instruments do meet high standards of reliability and validity.[21]

Observational assessment findings for individual teachers must be combined and summarized for all graduates of a specific program in order to draw conclusions about their program—or, perhaps, from large enough samples of program

graduates to produce reliable findings. Some programs are doing this already on their own and using the findings for assessing student teachers or new graduates. A few states even require this as part of their program approval system. Faculty use the same data for program improvement.[22] Finally, classroom assessment systems must have sufficient quality to be used reliably and effectively.

Two large national studies of reliable and valid classroom observation instruments now underway can help move us forward in using classroom assessment as an accountability measure. The Understanding Teacher Quality initiative (see http://www.utqstudy.org/index.html) is looking at the relationship between six instruments and pupil learning gains through videotaped observations of 450 teachers. Similarly, the Measuring Effective Teaching or MET effort is doing the same work with 3,700 teachers in six school districts (see http://metproject.org/project).

Using classroom assessment for large-scale accountability has its challenges, but states and programs that want to understand how teachers produce positive outcomes for students will support investment in continued development and use of these observational strategies. States can start now to support efforts to make the empirical links between classroom observation and student achievement—and they can organize training for assessors so that findings are reliable.

Persistence rates in teaching of program graduates should be reported for every teacher preparation program

Public disclosure of this information for up to five years postcompletion will stimulate progress in addressing high teacher turnover rates by drawing attention from teacher education programs, schools, districts, and policymakers.

Teacher turnover rates have drawn significant attention in recent years, and the disproportionate impact of teacher churn on particular kinds of schools and students is well documented. The National Commission on Teaching and America's Future argued in 2003 that teacher turnover undermines schools' capacity to initiate and sustain systemic improvement.[23] And a 2009 Consortium on Chicago School Research study further notes that, "High turnover rates produce a range of organizational problems for schools … thwart efforts to develop a professional learning community among teachers and make it difficult to develop sustained partnerships with the local community."[24]

Teacher turnover is a particular problem in low-achieving schools that have high proportions of students in poverty and minority students. Studies of student learning show that teachers become more effective (up to a point) as they gain experience, so teacher turnover also undercuts student academic achievement. But it's not just students who suffer. Preparation programs, schools, districts, and students all have a stake in a quality teacher workforce, and all of them are affected by a systemic problem like teacher turnover.

High rates of turnover persist despite the fact that many teacher education programs say they prepare teachers for challenging schools in urban or rural settings. To be fair, preparation programs are not solely responsible for turnover or its solution, but many programs don't even know if their graduates end up in classrooms, much less how long they stay in the profession. Many also aren't sure whether their graduates teach in the kinds of schools the program believes it is preparing them for.

Given teacher turnover's causes and consequences we need to align producers and employers through incentives, rewards, and better public information about the problem. A major step in this direction is to make sure that all the producers of new teachers disclose the teaching persistence rates of their graduates annually for at least five years after program completion.[25] States should report these persistence rates publicly for every program.

K-12 schools are already held accountable for teacher turnover. High rates of turnover result in weaker student academic learning gains than would otherwise be the case. Moreover, several states already use teacher survey data on school working conditions to push improvements at the building level, and they hold school leaders accountable for poor working conditions. The Obama administration has proposed reauthorization policies for the Elementary and Secondary Education Act that include teacher survey findings of school working conditions as part of the state reporting system for K-12 education.[26]

Studies show that preparation matters when it comes to teacher effectiveness.[27] It's particularly important where candidates obtain their clinical experience before graduation—in carefully selected and well-supported clinical sites—and how the program's clinical component is organized and supported by faculty so that graduates develop the skills and abilities needed in schools. Strong teachers are more likely to stay in the field.

Since teacher preparation programs already interact with schools they should help solve the turnover problem. But this can only happen through stronger incentives, including full public disclosure by the states of program-specific teacher persistence rates. Public disclosure of this information as a program quality indicator would give incentives for programs to pay more attention to teacher preparation for high-need schools.

Putting it into practice

Some teacher preparation programs study the persistence rates of their own graduates. But a reliable system for state accountability will need data systems that enable all programs to locate their graduates in the schools and districts where they teach.[28] Rather than have individual programs gather and publish these persistence rates, it might be more feasible for states to collect and disseminate the persistence rates for every program.

Since teacher preparation programs already interact with schools they should help solve the turnover problem.

Many programs produce a small number of graduates each year, and though persistence rate calculations are equally valuable for these programs, the need for solid inferences drawn from a small number of graduates suggests that pooling results across several years for small programs would help to offset any distortions caused by a small number of cases in a single cohort of program graduates.

Another issue in tracking teacher persistence is teachers who "stop out" after a few years of teaching for personal reasons and then return to the field. The Illinois Education Research Council found that about one-third of teachers who left teaching in the first few years of their careers later returned to teaching.[29] Annual teacher persistence reports that build a five-year cumulative record for a program cohort—for example, those who finish in 2010—would help solve this problem. A second or third year "dip" in persistence for a particular group of graduates would be offset in years four and five for those who take a short break and then return to teaching.

Finally, it's worth saying again that teacher education programs should not bear sole responsibility for teacher turnover. Effective responses to a serious issue in education require all players in the system to take responsibility for developing, training, and supporting teachers. Schools can't fix the turnover problem in isolation. The human capital "supply chain" has to be part of the solution.

Feedback surveys from preparation program graduates and from their employers should be part of state program accountability

A growing number of teacher education programs seek regular feedback from their graduates about the program and how well it prepared them to teach. Some programs solicit similar feedback from their graduates' employers. These survey findings let programs know in specific detail how well the graduates believe they were prepared for classroom teaching. And they allow employers to tell the program how they rate the graduates who were hired in their school or district: Were they ready to be successful beginning teachers? Would they hire program graduates in the future?

Nearly everyone who talks to schools or districts about their new teachers hears anecdotes from district human resources officials or from school principals about the graduates of this or that program. Some report they are so happy with a particular program's graduates that they can't get enough of them and would hire every graduate from that program if they could. Others are less positive, and some even say they would never hire a graduate from such-and-such program. But whatever we may read into these stories they do not provide systematic feedback about program or individual teacher quality. Surveys and response rates must meet standards of quality for feedback data to be reliable.

Those who have employed systematic feedback surveys know they can be very useful. Positive feedback lets program faculty know there's a good fit between the program and the schools it serves. Critical responses flag areas of concern that require the faculty to improve or the state to intervene. And while this information is not strong enough to stand on its own as an accountability measure, it still can be valuable as an indicator of program quality. Feedback survey data from graduates and their employers adds to the overall picture of program performance if it's used in concert with rigorous data on pupil learning gains, information on classroom teaching skills, and teaching persistence rates.

In addition to surveys individual programs conduct, there are good examples of multiprogram or statewide feedback surveys that can serve as models for building improved state accountability systems. The California State University System has conducted regular surveys of all teacher education program graduates and their employers since 2001, using a common instrument (see http://tinyurl.com/yetuw85). And the Pathways research project in New York developed preparation

program graduate follow-up surveys that it administered from 2004 to 2006. The project obtained responses from the graduates of several dozen programs (see http://tinyurl.com/ybgufex). Survey instruments and survey results are accessible online in each case.

The Consortium on Chicago School Research has surveyed teachers from multiple programs employed in Chicago public school classrooms (http://tinyurl.com/yeabgel), and the Woodrow Wilson National Fellowship Foundation recently reported the results of a survey of more than 500 "career changers" from many different programs (http://tinyurl.com/yeaf4qh).

Conducting feedback surveys is a practical and suitable strategy for state accountability.

In the cases mentioned here—state university system, complex research projects, state program improvement initiative, or a foundation interested in expanding the pool of high-quality teachers in the United States—the surveys of teachers and employers are not ends in themselves. They were envisioned as sources of useful information to support accountability and program improvement tools. So far, however, they are not regular tools of state accountability and program improvement.[30]

Putting it into practice

Conducting feedback surveys is a practical and suitable strategy for state accountability. But survey response rates at the program level have to be high enough for the programs themselves and for state regulators to have confidence in the results. Improved state data systems now coming online make it easier for programs and states to recruit survey respondents by linking practicing teachers back to the programs that produced them.

Countless universities and even high schools now tap the resources of the National Student Clearinghouse, the National Opinion Research Center, Harris Connect, and other organizations to obtain accurate up-to-date information on the whereabouts and work status of their graduates. This can be helpful to the many programs whose graduates are employed as new teachers in states other than the one where the program is located. Ratcheting up the importance and visibility of a measure will create powerful incentives for preparation programs and state regulators to devote the necessary resources to develop strong survey instruments with acceptable response rates—as with the persistence indicator discussed earlier.

Furthermore, programs should be expected to conduct surveys of graduates and employers annually, or no less than every other year, for the feedback survey indicator to be meaningful. Less frequent surveys will yield information too "stale" for

use as a program improvement tool. States also should make the feedback survey findings public since accountability is a public responsibility. Survey results by individual programs ought to be transparently understandable to customers and would-be customers.

Without taking business world analogies too far, districts, schools, and students are "customers" of the preparation programs. Customer views about quality and quality improvement are valuable pieces of information for programs, and feedback survey findings will help state officials charged with the responsibility to ensure that every child has an effective teacher.

A new system of teacher tests should be designed and implemented for state accountability as an indicator of program quality

The number of teacher tests now in use should be cut by more than 90 percent to make teacher testing an effective indicator of individual and program quality. And every state should adopt the same tests and the same pass rate policies.

States should immediately hold teacher education programs accountable through a limited number of common teacher tests. Tests could still be given at different stages of preparation: Content knowledge or pedagogical content knowledge tests might be employed before student teaching or as a requirement for program completion, with pass rates disclosed for all test-takers, not just for those allowed to complete the program.

States should immediately hold teacher education programs accountable through a limited number of common teacher tests.

At the same time, better tests and higher passing scores should be adopted. Reliable and valid measures of content knowledge and professional knowledge can be valuable to assess the quality of individual teachers and the programs that produced them.[31]

The five features of a new teacher testing system below would ensure that testing becomes a meaningful component of strong state accountability for teacher education programs:

- A vastly reduced number of tests developed with predictive validity in mind
- Elimination of basic skills tests, with rigorous program admissions standards used to weed out the weakest students before they enter teaching preparation programs

- New rigorous tests of subject area content knowledge and professional knowledge, pegged to meaningful levels of knowledge and performance by grade levels and subject areas

- Cut scores for passing the tests established at levels high enough to ensure that only the strongest prospective teachers are allowed to complete preparation programs and obtain an initial license to teach in public school classrooms

- Identical tests and passing scores adopted in every state

A variety of studies clearly show that teachers do learn professional knowledge and skills from their preparation programs, but there is "little evidence about the degree to which the learned skills contribute to teacher effectiveness."[32] This weak linkage may owe more to poor measures of teacher knowledge and skills than anything else. Revamped and high-quality teacher tests would bring greater rigor to assessing program and candidate quality, highlighting the factors that contribute to successful teacher preparation and making it more likely that measures of teacher knowledge and skills have predictive validity for pupil learning outcomes.

Current teacher testing fails as a quality control measure

Teacher testing per se is not responsible for the failure of current accountability policies, as noted earlier. Credible and effective accountability is instead undercut by the current tests themselves and how they're used. Many of them measure eighth-grade levels of knowledge, and cut scores for passing these tests are set low enough to guarantee that nearly every graduate will pass. In fact, 96 percent of program completers in the United States passed all state tests in 2006, as mentioned earlier.[33]

Teachers, schools, and students need a set of teacher tests that better predict classroom teaching performance and are equally good at predicting positive learning outcomes for K-12 pupils. Clear signals about program and graduate quality get lost in the noise of so many tests, test categories, and pass rate policies. The current structure serves the interests of testing companies, state regulators trying to avoid real accountability, and their allies in preparation programs. It does nothing to help those programs that want to produce effective graduates, and it doesn't provide a useful tool for state oversight of program quality.

States should therefore work with the testing companies and with teacher educators committed to quality to develop and implement a new system of teacher testing. Instead of the current system with more than 1,100 different teacher tests, a new system of higher quality and more relevant tests could be organized around the matrix of test categories and grade levels below.

A new system of teacher testing

Matrix below could simplify the current system of 1,100 different teacher tests

Test category	Grades Pre K-2	Grades 3-5	Grades 6-8	Grades 9-12	Special education
Content knowledge					
Professional knowledge					

Based on teaching assignments and K-12 subject areas at the various grade levels, content areas for testing would include reading, English language arts, math, social studies, biology, chemistry, earth science, and physics, as well as areas such as the arts, computer science-technology, and English as a second language. Special education teachers should also be tested in content areas as well as in their specialized knowledge.

The grade configuration proposed here may not fit the licensure and certification structure of every state—another byproduct of the idiosyncratic local regulation of a national profession—and other clusters of grades may make sense. But the key point of this example is to significantly reduce the number of teacher tests.

The proposed array illustrated above results in two test categories, 11 content areas, and four grade levels, including special education. A redesigned system of teacher testing would thus include about 90 different tests—instead of 1,100—covering all grades and subject areas as well as two key forms of teacher knowledge (subject area content and professional knowledge).

Some currently used teacher tests purport to measure teaching skills, but they do not have predictive validity and are not tied to pupil outcomes. The state accountability system proposed here will capture reliable and valid information about actual classroom teaching performance (see the section on using program graduates' classroom performance to judge the quality of all teacher preparation programs).

Notably, this proposed structure for teacher tests in the new state accountability system does not include any tests in the "basic skills" category that states currently use. As described earlier, these are essentially eighth-grade level assessments of basic content knowledge, and a robust system of accountability for programs and graduates must aim much higher than eighth-grade knowledge for teachers and teacher candidates. Program-level selection processes should ensure that only academically able college students are allowed to enroll in a teacher education program. It was (and is) the failure of many programs to have adequate admissions standards in the first place that led to "basic skills" testing by the states.

Full deployment of the accountability measures described in this paper will make it impossible for the weakest college students to make it through preparation programs. The revamped accountability system will eliminate the need for "basic skills" tests by invoking real consequences for weak programs.[34]

Every state adopts common policies and program quality standards

Every state needs to adopt the new system proposed here for it to be most effective. Other professions have embraced accountability policies and practices for programs and graduates that are the same in all 50 states. One result of this is that quality has the same definition no matter where the program is or where the graduate works.

In contrast, there are more than 50 versions of quality standards, policies, and accountability systems for teaching and teacher preparation. This is a major weakness of the current system. The proposed new system replaces a weak and largely irrelevant system of state teacher education program accountability with a set of consistent indicators that assess important aspects of program and teacher quality.

There are four compelling reasons why all states should embrace the teacher preparation program accountability system described here.

Teachers prepared and licensed in different states: A surprising number of newly licensed teachers in the United States complete a preparation program in one state and obtain their initial license somewhere else. According to Education Secretary Arne Duncan's most recent teacher quality report to Congress, 20 percent of initial licenses in 32 states were granted to new teachers whose preparation for teaching took place in another state. For another 12 states and the District of Columbia 40 percent of initially certified teachers were prepared by a teacher education program in another state.[35]

There are understandable geographic and personal reasons for these patterns. But the fact is that students and schools in dozens of states are at the mercy of program quality policies from some other state over which they have no influence. Since subject-area knowledge and effective teaching skills are not bound by geography, the rules by which programs and graduates are judged on this knowledge and skills should not be constrained by state lines.

Student mobility across state lines: The knowledge and skills K-12 students need to be successful in school, work, and life are not state specific. Yet the absence of more uniformity in accountability measures means that students who move from one state to another are likely to encounter public school teachers whose preparation programs were subject to different quality yardsticks. Fortunately, educators and policymakers are beginning to understand the value of student learning

standards that are the same in every state (see next point). Teacher quality rules should transcend state lines and be the same across the country.

Emerging common standards and assessments: The Common Core State Standards Initiative (http://www.corestandards.org/) requires stronger and more universal policies for preparation program accountability and teacher quality than is possible under the current system. It is supported to date by 48 states. A single set of accountability policies and practices in every state can be embraced by preparation programs and regulators as indicators of quality and assurances to the public.

Other professions and state accountability: States can also draw on the experience of other professions. Engineering, accountancy, nursing, and medicine operate with uniform state accountability standards and requirements. This has occurred without doing violence to professional autonomy or academic freedom among program faculty. Perhaps the most important point here is that all states have implemented a single set of accountability policies and practices without infringing on the principles of federalism.

Teacher quality rules should transcend state lines and be the same across the country.

In nursing, for instance, the NCLEX-RN is accepted by every state as the single licensure test that determines whether a program graduate is granted a license to practice nursing. Every state uses the same passing standard, and pass rates are tied to program accountability for more than 1,200 professional nursing programs in the United States (see https://www.ncsbn.org/nclex.htm).

Engineering has a similar story. All states employ the same battery of tests for would-be engineers, and every state employs the same passing score (see http://www.ncees.org/Exams.php).

Medical licensure standards in the United States can be summarized in one chart (see http://www.fsmb.org/usmle_eliinitial.html) because there is agreement across the states and within the profession about entry standards into the profession and quality standards for medical preparation programs. Accountancy follows a similar pattern with all states using the same four-part Uniform CPA Examination and passing scores (see http://www.bls.gov/oco/ocos001.htm#training).

Teachers and teacher education have many similarities to nursing and nurse preparation, so state policymakers may find nursing program oversight practices particularly relevant to teacher education. Like teaching, nursing is a predominantly

female profession with multiple preparation pathways—hospitals, community colleges, and universities, for example—and more than 1,200 different nursing education providers exist. The academic quality of entrants into the nursing profession has parallels to teaching.

Nursing programs and graduates are subject to the licensing and program approval authority of the states. In sharp contrast to teacher education, however, it is noteworthy that the nursing profession and the regulators support and have implemented the same set of oversight rules in every state.

The essentially national character of standards and practices for nursing and the other professions discussed in this section serves as a form of quality control that does not exist in teacher education. And the high degree of consensus within each field about the values, standards, and practices that should characterize professional preparation is a mechanism that links accreditation, state oversight, and professional licensure. This strategy protects the public with the same set of rules in every state, and it brings higher levels of public respect for the profession as a whole and for those who serve the public through their professional work. A strong uniform accountability structure for an entire profession is a demonstrably effective strategy for achieving secure professional status in the eyes of the American people.

Recommendations for states

Some of the changes proposed in this paper will take time—especially the development of high-quality tests for teacher candidates and new teachers. Even so, states can take significant steps now toward more rigorous accountability policies for teacher education programs. They can implement these four accountability measures with data systems that are already in place or on the horizon:

States can take significant steps now toward more rigorous accountability policies for teacher education programs.

- Tie K-12 pupil learning outcomes to preparation program graduates and hold the programs accountable for teacher effectiveness
- Begin to implement high-quality observational assessments of classroom teaching by supporting efforts to link these assessments to student achievement and by developing rigorous training for classroom observers to ensure reliable assessment findings
- Employ current state data systems to track the teaching persistence rates for graduates of every program, and use the findings as a public disclosure measure
- Implement feedback surveys of preparation program graduates and their employers using state education, labor department (or state insurance department), university, and school district data systems

Individual states can take these steps right away. Another option is for consortia of states to work together and implement identical accountability measures and performance criteria—just as groups of states are now working on common K-12 student assessments.

And finally, individual states, as well as state consortia, can move immediately to end the fiction that current teacher testing policies provide meaningful accountability. All states should raise passing cut-off scores on every test, at least to the

national average for all test-takers. States also should make dramatic reductions in the number of redundant tests used for licensure. But these two steps are only incremental changes on the way to the far more robust, simplified, and relevant testing system proposed in this paper for teacher candidates and new teachers, a new structure of accountability that should be adopted as the norm in every state.

Conclusion

Current state accountability policies for teacher education are not up to the task of ensuring that all programs routinely produce effective teachers for our public schools. Key outcomes that matter—including student achievement and classroom teaching performance, among others—are, with few exceptions, not used as quality control mechanisms by the states. Further, the record is clear that states seldom take action against weak programs.

Aside from these serious problems, our current fractured system of quality control is an obstacle to professional status for preparation programs and their graduates. Every state has its own set of policies and practices despite the fact that many students and teachers move regularly from one state to another. State-specific program accountability practices ignore the fundamentals of teaching and learning, but subject-area knowledge and teaching skills are universals unconstrained by state boundaries. Oversight for preparation program quality should function the same way.

A new accountability system for teacher education should be undergirded by empirically based performance and feedback indicators directly related to teaching and learning outcomes that affect students, teachers, and schools. Every state should use the same system of indicators, and meaningful quality control for teacher preparation programs requires the public and policymakers to have easy access to measures of quality.

Many teacher educators and program leaders understand that the current system is a failure because it focuses on irrelevant inputs and insulates weak programs from pressures to change or close down. The system now in place also undermines respect for the teaching profession and for teacher education as a form of professional education. A more powerful accountability system will send the right signals about program quality to professional educators, policymakers, schools, and parents.

References

Boyd, Donald, and others. 2009. "Teacher Preparation and Student Achievement." *Educational Evaluation and Policy Analysis* 31 (4): 416-440.

Clotfelter, Charles, Helen Ladd, and Jacob Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* 41 (4): 778-820.

Clotfelter, Charles, Helen Ladd, and Jacob Vigdor. 2007. "How and Why Do Teacher Credentials Matter for Student Achievement?" Working Paper 12828. National Bureau of Economic Research.

Crowe, Edward. 2008. "Teaching as a Profession: A Bridge Too Far?" In M. Cochran-Smith, S. Feiman-Nemser, D.J. McIntyre, and K. Demers, eds., *Handbook of Research on Teacher Education*. New York: Routledge, Taylor & Francis.

Goldhaber, Dan. 2006. "Everyone's Doing It, But What Does Teacher Testing Tell Us About Teacher Effectiveness?" Seattle: University of Washington Center for the Reinvention of Public Education.

Goldhaber, Dan, and Jane Hannaway, eds. 2010. "Creating a New Teaching Profession". Washington: The Urban Institute Press.

Harris, Douglas, and Tim Sass. 2006. "The Effects of Teacher Training on Teacher Value Added." Working Paper. Florida State University, Department of Economics.

National Academy of Education. 2009. "Teacher Quality: Education Policy White Paper." (Retrieved from the National Academy of Education, www.naeducation.org.)

Pianta, Robert, and Bridget Hamre. 2009. "Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity." *Educational Researcher* 38 (2): 109-119.

Rice, Jennifer King. 2003. *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*. Washington: Economic Policy Institute.

U.S. Department of Education, Office of Postsecondary Education. 2009. "The Secretary's Sixth Annual Report on Teacher Quality: A Highly Qualified Teacher in Every Classroom" Washington.

Endnotes

1 Suzanne Wilson and Peter Youngs, "Research on Accountability Processes in Teacher Education." In M. Cochran-Smith and K. Zeichner, eds., *Studying Teacher Education*. (Washington: AERA, 2005, p. 591-644).

2 See Crowe, "Teaching as a Profession."; Wilson and Youngs, "Research on Accountability Processes in Teacher Education."

3 Information on the use of pass rates by the states is from U.S. Department of Education, "The Secretary's Sixth Annual Report on Teacher Quality: A Highly Qualified Teacher in Every Classroom" (2009), p. 14.

4 The National Research Council quote is from page 47 of Karen J. Mitchell and others, *Testing Teacher Candidates: The Role of Licensure Tests in Improving Teacher Quality*. (Washington: National Academy of Sciences, 2001). The Wilson and Youngs discussion can be found on page 592 of the analysis cited above.

5 The absence of a direct connection between teacher licensure tests and classroom practice was a major theme of the 2001 NRC study, *Testing Teacher Candidates*, and the quote can be found on page 168 of this volume. Furthermore, other studies report that, at best, certification exam scores have modest impact on pupil learning and perhaps none at all. For more information, see Charles Clotfelter, Helen Ladd, and Jacob Vigdor, "Teacher-Student Matching and the Assessment of Teacher Effectiveness" Working Paper 11936. (Cambridge: National Bureau of Economic Research, 2006), as well as Charles Clotfelter, Helen Ladd, and Jacob Vigdor, "How and Why Do Teacher Credentials Matter for Student Achievement?" Working Paper 12828 (Cambridge: National Bureau of Economic Research, 2007); and Douglas Harris and Tim Sass, "The Effects of Teacher Training on Teacher Value Added" Working Paper (Florida State University, Department of Economics, 2006). For another perspective, see Dan Goldhaber, "Everyone's Doing It, But What Does Teacher Testing Tell Us About Teacher Effectiveness?" (Seattle: University of Washington Center for the Reinvention of Public Education, 2006).

6 U.S. Department of Education, "The Secretary's Sixth Annual Report on Teacher Quality: A Highly Qualified Teacher in Every Classroom" (2009), p. 36.

7 See ScienceDaily, "US Needs Better-Trained Math Teachers to Compete Globally, Study Finds," Press release, April 19, 2010, available at http://www.sciencedaily.com/releases/2010/04/100415105954.htm; and the full study, Michigan State University, College of Education, "Breaking the Cycle: An International Comparison of U.S. Mathematics Teacher Preparation" available at http://tinyurl.com/37wyql6.

8 Descriptions of the current testing scheme for teacher candidates and new teachers can be found in U.S. Department of Education, "The Secretary's Sixth Annual Report on Teacher Quality: A Highly Qualified Teacher in Every Classroom" (2009). Figures for the number of tests in each testing category come from page 23 and the single-state example is found on page 25.

9 "Important teaching knowledge" includes the content knowledge, pedagogical content knowledge, and related teaching skills that form the backbone for classroom observation instruments such as CLASS, MQI (Mathematical Quality of Instruction), and the Intellectual Demand Assignment Protocol (IDAQ). For more see http://www.utqstudy.org/instruments.html.

10 Sources are U.S. Department of Education, "The Secretary's Sixth Annual Report on Teacher Quality: A Highly Qualified Teacher in Every Classroom" (2009), p. 12-13; and Section 208 of the HEA Title II statute.

11 Wilson and Youngs, "Research on Accountability Processes in Teacher Education," p. 593.

12 See The Wallace Foundation, "Assessing the Effectiveness of School Leaders: New Directions and New Processes" (2009), p. 5, available at http://tinyurl.com/ydv6ltc.

13 For more information on Louisiana's system as well as the policies and research behind its development, see State of Louisiana, Board of Regents, "Teacher Education Initiatives," available at http://tinyurl.com/27y5fzg.

14 Details on the Texas system come from state senate bill 174 and chapter 229 of the Texas Administrative Code, both adopted in 2009. Information on plans in Delaware and Tennessee can be found in their successful RTT proposals (see U.S. Department of Education, "Race to the Top Fund: States' Applications, Scores and Comments for Phase 1," (2010), available at http://tinyurl.com/yk5skfg).

15 Laura Goe, Courtney Bell, and Olivia Little, "Approaches to Evaluating Teacher Effectiveness: A Research Synthesis" (Washington: National Comprehensive Center for Teacher Quality, 2008), p. 47.

16 As noted, the challenges include data systems, the tests themselves, and the methodologies. Progress on data systems and tests—especially the emerging consortium that includes the U.S. Department of Education, states, and some education funders—with possible multistate adoption of student assessments linked to multistate adoption of the Common Core State Standards Initiative shows that people of good will are working to address the practical challenges of robust accountability systems. We know that current program quality indicators are ineffective and often misused to protect weak programs. Student learning is a critical outcome. Working together, states, the federal government, foundations, and the profession can implement and make effective use of this measure of teacher effectiveness—as we see already in Louisiana.

17 Goe, Bell, and Little, "Approaches to Evaluating Teacher Effectiveness," p. 9.

18 The importance of new teacher performance—and the relative unimportance of degrees or credentials—is discussed by Jonah Rockoff and others, "Can You Recognize an Effective Teacher When You Recruit One?" Working Paper 14485 (Cambridge: National Bureau of Economic Research, 2008), p. 3-4. The strong quality of some current classroom observation strategies is discussed in Robert Pianta and Bridget Hamre, "Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity" *Educational Researcher* 38 (2) (2009): 109-119.

19 Goe, Bell, and Little, "Approaches to Evaluating Teacher Effectiveness," p. 20.

20 Pianta and Hamre, "Conceptualization, Measurement, and Improvement of Classroom Processes," p. 109.

21 Excellent discussions of these issues can be found in Pianta and Hamre, "Conceptualization, Measurement, and Improvement of Classroom Processes," p. 111; and in Goe, Bell, and Little, "Approaches to Evaluating Teacher Effectiveness," p. 22.

22 To be useful and relevant for program accountability purposes, assessment of classroom teaching must capture information about teaching practices that are associated with improved pupil learning outcomes. This is now possible: a variety of carefully developed observation instruments have been shown empirically to predict student-learning gains (Pianta and Hamre, "Conceptualization, Measurement, and Improvement of Classroom Processes," p. 115). Other studies of currently available classroom teaching assessments also report positive relationships between teacher scores on the assessment and gains in student achievement (see Goe, Bell, and Little, "Approaches to Evaluating Teacher Effectiveness," p. 22; Anthony Milanowski, "The Relationship Between Teacher Performance Evaluation Scores and Student Achievement: The Evidence from Cincinnati," *Peabody Journal of Education* 79 (4) (2004): 33-53; as well as A. Milanowski, Steven Kimball, and Allan Odden, "Teacher Accountability Measures and Links to Learning." In L. Stiefel, A. Schwartz, R. Rubinstein, and J. Zabel, eds., *Measuring School Performance and Efficiency*. (Larchmont: Eye on Education, 2005).

23 Studies and reports on teacher turnover include work by NCTAF in *No Dream Denied* and their 2007 study of teacher turnover in five school districts (see National Commission on Teaching and America's Future, "The Cost of Teacher Turnover Study and Cost Calculator" (2007), available at http://tinyurl.com/22wasx), work by Smith and Ingersoll (2004), and the study of turnover in Illinois by White et al. (2008). More recently, the Consortium on Chicago School Research provided a very detailed analysis of teacher turnover and its impact of particular schools and students. See Elaine Allensworth, Stephen Ponisciak, and Christopher Mazzeo, "The Schools Teachers Leave: Teacher Mobility in Chicago Public Schools" (Chicago: Consortium on Chicago School Research, University of Chicago, 2009).

24 Allensworth, Ponisciak, and Mazzeo, "The Schools Teachers Leave," p. 3.

25 See the recommendations of the National Council on Teacher Quality (NCTQ) in their 2010 report on state policies regarding teacher quality: National Council on Teacher Quality, "2009 State Teacher Quality Yearbook" (2009), available at http://www.nctq.org/stpy09/.

26 At the same time, of course, teacher feedback surveys (like the school working conditions surveys used in North Carolina, Colorado, and a few other states where thousands of teachers provide useful feedback on the schools where they teach) and analyses of pupil learning outcomes can help policymakers and programs understand the context for program-wide persistence rates, whether they are high, low, or in-between. Examples include Colorado (http://tellcolorado.org/) as well as North Carolina (http://www.ncteachingconditions.org/).

27 See Boyd, Donald, and others, "Teacher Preparation and Student Achievement."; Harris, Douglas, and Sass, "Teacher Training, Teacher Quality and Student Achievement."; and the essays in Dan Goldhaber and Jane Hannaway, eds., "Creating a New Teaching Profession" (Washington: The Urban Institute Press, 2010).

28 These systems are becoming increasingly common in the states, in part thanks to support from the State Longitudinal Data Systems (SLDS) program of the U.S. Department of Education. A 2003 report from SHEEO describes parameters needed for an effective system. Even where states prohibit or discourage data system linkages between teachers and their pupils, teacher and teacher employment data is becoming more accessible each year. The SLDS program provides grants to states to improve the quality of their state education data systems. See http://nces.ed.gov/Programs/SLDS/resources.asp. The State Higher Education Executive Officers (SHEEO) and the Education Commission of the States (ECS) published an excellent guide to development of state data systems that capture the full range of information needed to make and evaluate polices regarding teacher quality. See Richard Voorhees and Gary Barnes, with Richard Rothman, "Data Systems to Enhance Teacher Quality" (Denver: State Higher Education Executive Officers, 2003).

29 Karen DeAngelis and Jennifer Presley, "Leaving Schools or Leaving the Profession: Setting Illinois's Record Straight on New Teacher Attrition" (Springfield: Illinois Education Research Council, 2007). They found that while 67 percent of teachers in Illinois leave their first teaching assignment within five years, only 27 percent leave the profession of teaching in the state over the same period of time, because about one-third of the initial leavers return to teaching in the state (see p. 3).

30 As noted earlier, it appears as though Texas will implement feedback surveys from school principals who hire program graduates, as well as exit surveys, data on persistence of graduates in teaching, and analyses of pupil learning outcomes. Statutory and administrative language has been published, but implementation procedures are not yet developed.

31 Mitchell and others, *Testing Teacher Candidates: The Role of Licensure Tests in Improving Teacher Quality*, p. 166; Wilson and Youngs, "Research on Accountability Processes in Teacher Education." p. 617.

32 Jennifer King Rice, *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*. (Washington: Economic Policy Institute, 2003), p. 28.

33 Furthermore, the states and testing companies operate a massive array of "basic skills," "content knowledge," and "teaching skills" assessments: more than a thousand *different* tests, including about 160 "basic skills" tests, more than 100 different tests of "professional knowledge", and over *800* different content knowledge tests. On the K-12 grade level of teacher tests, see Mitchell and others, *Testing Teacher Candidates: The Role of Licensure Tests in Improving Teacher Quality*, p. 51. National cumulative pass rates are from U.S. Department of Education, "The Secretary's Sixth Annual Report on Teacher Quality." p. 36; test category labels and counts of tests within each category are from page 23 of the same source.

34 Teacher tests currently in use by the state are classified into various categories by the federal Title II reporting system: basic skills, professional knowledge, academic content, other content, teaching special populations, and "performance" tests. There are about 1700 tests in these categories. See U.S. Department of Education, "The Secretary's Sixth Annual Report on Teacher Quality," p. 23.

35 Ibid, p. 44.

About the author

Edward Crowe is a consultant on teacher quality policy for several organizations and projects, including the Carnegie Corporation of New York and the Woodrow Wilson National Fellowship Foundation. He previously worked as senior consultant for the National Commission on Teaching and America's Future, where he focused on teacher preparation projects, research on the cost of teacher turnover, and efforts to improve identification and development of high achieving students for careers in science, mathematics, engineering and technology disciplines. He also served as an adviser to the Hunter Foundation of Scotland and the Scottish National Executive on teacher quality, and currently serves on the Advisory Council for the Texas Center for Research, Evaluation and Advancement of Teacher Education. Crowe also has worked with the New York City Partnership for Teacher Excellence and with the Committee on Teacher Preparation Programs for the National Research Council.

Crowe was the first director of the Title II Teacher Quality Enhancement Program for the U.S. Department of Education. He has experience in state higher education policy as well, having worked for the University of North Carolina system, with the Arkansas state higher education coordinating board, and as senior manager of a National Science Foundation-funded statewide math and science education reform project.

He is a graduate of Boston College, and holds master's and doctoral degrees in political science from The University of North Carolina-Chapel Hill.

Acknowledgements

The author expresses thanks to several people for their ideas and assistance with this project: Robin Chait, Center for American Progress; Rena Subotnik, American Psychological Association; Joanne Bogart, U.S. Department of Education; and Michael Allen, Allen Education. For expert review and very helpful suggestions, thanks to Jeanne Burns, Louisiana Board of Regents; Heather Harding, Teach for America; and Sandi Jacobs, National Council on Teacher Quality.

The Center for American Progress is a nonpartisan research and educational institute dedicated to promoting a strong, just and free America that ensures opportunity for all. We believe that Americans are bound together by a common commitment to these values and we aspire to ensure that our national policies reflect these values. We work to find progressive and pragmatic solutions to significant domestic and international problems and develop policy proposals that foster a government that is "of the people, by the people, and for the people."

Center for American Progress

1333 H STREET, NW, 10TH FLOOR, WASHINGTON, DC 20005 • TEL: 202-682-1611 • FAX: 202-682-1867 • WWW.AMERICANPROGRESS.ORG